(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 February 2002 (07.02.2002)

PCT

(10) International Publication Number
WO 02/11446 A2

(51) International Patent Classification[7]: H04N 7/173

(21) International Application Number: PCT/EP01/07965

(22) International Filing Date: 11 July 2001 (11.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/627,188 27 July 2000 (27.07.2000) US

(71) Applicant: **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors: **MCGEE, Thomas**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **DIMITROVA, Nevenka**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **AGNIHOTRI, Lalitha**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agent: **GROENENDAAL, Antonius, W., M.**; Internationaal Octrooibureau B.V., Prof Holstlaan 6, NL-5656 AA Eindhoven (NL).

(81) Designated States *(national)*: CN, JP, KR.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/11446 A2

(54) Title: TRANSCRIPT TRIGGERS FOR VIDEO ENHANCEMENT

(57) Abstract: A system and method for retrieving information supplemental to video programming. Transcript text is searched for terms of interest and information associated with the terms is identified. Depending upon a user profile and the category of video segment being viewed, the supplemental information is formatted for display. Over time, the rules for associating the supplemental information with the terms of interest may be modified using a learning model.

Transcript triggers for video enhancement

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention is directed to the field of media technology. It is particularly directed to video and related transcript text.

2. Cross-Reference to Related Applications

This invention associates video with supplementary information using a text transcript, and extracts and augments textual features, as does co-pending application, Ser Nr. 09/351,086, Filed 1999 July 9 by the assignee, and incorporated by reference herein.

3. Description of the Related Art

In recent years, the number of media sources has increased and the volume of information from each source has also increased, resulting in information overload. Most consumers have neither the time nor the inclination to sift through the morass of information for what is pertinent to their wants and needs. Accordingly, so called "push technology" has developed. Webcasting applications such as Pointcast or Backweb, or the newer web browsers, ask the user which information categories and web sites the user is interested in. A web server then "pushes " information of interest to the user instead of waiting until the user requests it. This is done periodically and in an unobtrusive manner.

Concurrently, as media technology has progressed, the lines between video, audio, and other media have been blurred. Advances in media technology have enabled the delivery of Internet information and other informational material to the consumer's video display, along with the traditional television programming. Because the Internet has become a tool of e-commerce, consumers are conditioned to view a combination of media, video, audio, and text information on the same or associated topics. Consumers are acquainted with the hyperlink concept and the notion of "drilling down" to retrieve additional information on a subject they are viewing on the World Wide Web (WWW).

Retrieval of this additional information can currently be accomplished using closed caption text, audio, and automated story segmentation and identification. The Broadcast News Editor (BNE), provided by Mitre Corporation, enables such retrieval by automatically partitioning newscasts into individual story segments, and providing a

summary of each story segment in the first line of the closed-caption text associated with the segment. Keywords from the closed-caption text or audio are also determined for each story segment.

The Broadcast News Navigator (BNN), also from Mitre Corporation, sorts story segments by the number of keywords in each story segment that match search words selected by the consumer. Accordingly, story segments likely to be of interest to a particular consumer can be readily identified. However, using a combination of BNN and BNE requires that the consumer have an explicit search topic in mind, which is usually not the case in a typical channel-surfing scenario.

Patents which disclose providing the user with information supplemental to a television program include US Patent No. 5,809,471 to Brodsky entitled "Retrieval of additional information not found in interactive TV or telephony signal by application using dynamically extracted vocabulary" and US Patent No. 6,005,565 to Legall et al. entitled "Integrated search of electronic program guide, internet and other information resources." In the '471 patent, keywords are extracted from a television program or closed caption text, creating a dynamically changing dictionary. The user requests information based upon an item seen or word heard in the television broadcast. The user's request is matched against the dictionary, and when there is a match, a search for supplemental information to display is initiated.

In the '565 patent, the user selects topics and sources to search. Based on the user input, the search tool performs a search of the electronic program guide and other information resources such as the World Wide Web, and displays the results. Both the '471 patent and the '565 patent require that the user provide a keyword of interest. Neither patent relates the supplementary information retrieved to the global context of the program, (i.e. news program), as opposed to the subject matter of the program (i.e. the Stock Market report).

SUMMARY OF THE INVENTION

Accordingly, it would be advantageous to provide a method and system employing transcript text, for automatically providing supplementary multimedia information enhancing the consumer's television viewing experience. So called transcript text is comprised of at least one of the following: video text, text generated by speech recognition software, program transcripts, electronic program guide information, and closed caption text that contains all or part of the program information. Video text, is superimposed or overlaid

text displayed in the foreground, with the image as a background. Anchor names for example, often appear as video text. Video text may also take the form of embedded text, for example, a street sign that can be identified and extracted from the video image.

It would also be advantageous to provide supplementary information, which is specific not just to the individual consumer's known interests or profile, but also to the context of the program being viewed. For example, news segments would be associated with links to the Cable Network News (CNN) Web page while commercials would be associated with additional product information. The method and system would use learning models to continually develop new associations between the television content and other media content as well as to customize which type and how much supplementary information should be displayed. In this way, supplementary information would be integrated seamlessly with a television program without disturbing the viewer or requiring any action on the viewer's part.

The present invention addresses the foregoing needs by providing a system, (i.e., a method, an apparatus, and computer-executable process steps), for retrieval of supplementary information associated with a video segment, for display on the consumer's video display. The system includes a recognition engine for determining whether expanded keywords for retrieving supplementary information are contained in the closed captioned text accompanying the video segment or in other transcript related text. If a keyword is found, a stored rule indicates the supplementary information to be displayed, the information having been selected from a larger set of information, and selected in accordance with a user profile and the context of the segment. Alternatively, the transcript keywords are expanded and then matched to the user's profile. The context of the segment is automatically determined based upon classification data. These data include the program classification, object tracking methods, natural language processing of transcript information and/or electronic program guide information.

The information is displayed in a window or superimposed unobtrusively over the main video segment. Alternatively, the information is transmitted, for example to a hand-held device or an email account, stored to secondary storage, or cached in local memory. The system automatically recognizes the beginning and end of each segment, in the story classifications, and so is able to update the subset of rules to correspond to the program segment context.

In a further aspect of the invention, the set of rules for associating supplementary information with the video segment being viewed is dynamic and based upon a learning model. The set of rules is updated from a set of sources, including third-party

sources, and makes information available to the user in accordance with the user's choices and pattern of behavior. In one embodiment, the rules are transmitted from a Personal Digital Assistant (PDA) enabled with a wireless connection.

This brief summary has been provided so that the nature of the invention will be understood quickly. A more complete understanding of the invention is obtained by reference to the following detailed description of the preferred embodiments thereof in connection with the attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a system on which the present invention is implemented.

Figure 2 depicts elements of the processor contained within the system.

Figures 3a and 3b are flow diagrams used for explaining the operation of the present invention.

Figure 4 is a table illustrating supplementary information triggers for a given video segment, according to the present invention.

Figure 4a illustrates how keywords and triggers are expanded.

Figure 5 is a diagram of an embodiment of the invention illustrating a learning model.

Figure 6 is a diagram illustrating how the association rules database, for retrieving supplementary information, is updated and maintained.

Figure 7 is a diagram illustrating how supplementary information is displayed.

Figure 8 is a diagram illustrating one embodiment of the invention in which a set-top box is used.

Figure 9 is a diagram illustrating another embodiment of the invention in which a television display is used.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Figure 1 shows a representative embodiment of a system on which the present invention is implemented. In this embodiment, a multimedia processor system 6 includes a processor 12, a memory 10, input/output circuitry 8, and other circuitry and components well known to those skilled in the art. An analog video signal or a digital stream is input to the receiver 2. This stream is compliant with MPEG or other proprietary broadcast formats.

In accordance with the MPEG standard, video data is encoded using discrete cosine transform encoding and is arranged into variable length encoded data packets for

transmission. One version of the MPEG standard, MPEG-2 is described in the International Standards Organization — Moving Pictures Experts Group Document "Coding of Moving Pictures and Audio", ISO/IEC JTCI/SC29/WG11, July, 1996. MPEG is just one example of a format, which can be utilized in the system.

Transcript text, transmitted in the video signal 162, is extracted by the transcript extractor 4 from either line 21 of the analog video signal or the user data field of the MPEG stream. The transcript extractor 4 also partitions the video program into segments. The transcript text for the particular frame may be stored in the memory 10. Alternatively, it is analyzed as a real-time data stream.

Also stored in the memory 10 is Electronic Program Guide Information (EPG). This information, describing television broadcast information for a period of days or weeks, is downloaded on user request or at a preprogrammed time. It is transmitted by local analog TV broadcasters over the vertical blanking interval or through MPEG-2 private tables on a "home barker" channel. It can also be transmitted via telephone line or through wireless means. EPG data includes information such as the program's genre and subgenre, its rating, and a short program description. EPG data is used to determine the context of a program, such as whether it is a news program, a paid programming excerpt, a soap opera, or a travelogue.

Also stored in secondary storage 18 and available in the memory 10 is personal profile information, in the form of keywords or "triggers," describing the user's interests. Typical triggers could be "Clint Eastwood", "environment", "presidential election" or "hockey". These triggers are expanded in one aspect of the invention to include synonymous and related terms.

As is well known in the prior art, a personal profile of the user's interests is established automatically, by user input, or by a combination of both methods. For example, the TiVo™ Personal TV Service allows the user to indicate which programs the user prefers using a "Thumbs Up" or "Thumbs Down" button on the TiVo™ remote. TiVo™ then builds upon this information to select other related programs the user likes to view.

When a trigger matches keywords contained in the transcript text, supplementary data is retrieved, for example from the Internet 14 or proprietary sources 13 through the communication means 17. Another source for supplementary data is, for example, another channel. The data is then displayed to the user on a display 16 either as a Web page or a portion thereof or superimposed over the main video in a non-intrusive

fashion. Alternatively or additionally, a simple Uniform Resource Locator (URL) or informative message is returned to the viewer.

Rules for associating these triggers with supplementary data such as World Wide Web (WWW) pages are also stored in the secondary memory 18 and available from the memory 10. These rules are established through a default profile that is updated based on user behavior, or though a query program that prompts the user for interests and then generates the rule set. The rules are also received from a mobile device 15 such as a Personal Digital Assistant (PDA) or cell phone through the communications means 17. These rules associate supplementary information with the triggers, depending on the context of the program segment being viewed. For example, if a program segment is an advertisement for Clint Eastwood's new movie, the context is commercial and the supplementary data retrieved is a description of the movie he is starring in. If a program segment is a description of Clint Eastwood's car accident, the context is news, and the supplementary data retrieved is a biographical web page or a link to www.cnn.com to obtain more information about why he is in the news.

As illustrated above, association rules are also dependent upon a combination of EPG fields. For example, if "Clint Eastwood" appears in the actor's field of the EPG data, and the context is determined to be commercial, and the closed caption data is "We will be returning shortly to Clint Eastwood and Fist Full of Dollars after these announcements," then, the association rule retrieves supplementary data pertaining to the particular movie being shown. On the other hand, if "Clint Eastwood" does not appear in the actor's field of the EPG data, and the context is commercial, and the closed caption data is "High Plains Drifter starring Clint Eastwood will be aired on Friday," then, the association rule retrieves supplementary data pertaining to showtimes for the movie. These differences can be determined, for example, by comparing the text of the credits with text extracted from the closed caption data. It there is a match, then the program being advertised is the program being viewed. Alternatively, natural language processing can be used to identify key phrases such as "returning to" which would also indicate that the program being advertised is the program being viewed.

Alternatively, if "Clint Eastwood" does not appear in the actor's field of the EPG data, and the context is commercial, and the closed caption data says "Clint Eastwood's new movie will be released shortly", then the association rule retrieves supplementary data by linking to the Clint Eastwood home page to find out more about the movie.

Association rules also determine the category of media to be retrieved. For example, if "Kosovo" is the trigger and the program is sponsored by National Geographic, the association rule retrieves a map of the region. Alternatively, if the program segment context is news and the word "war" is located in the EPG data, then the association rule retrieves a recent political history of the region.

In alternative embodiments, the system includes a video display with built-in processing and memory, or a separate set top box for processing and storing information. These embodiments can include communication means or interface to communication means. Receipt of the video signal and Internet information is via wireless, satellite, cable or other media. This system is modifiable to transmit the supplementary information via the communication means 17 as an output signal over a radio transmitter, or via wireless means, where the signal is embodied in a carrier wave 160. The supplementary information is transmittable to an e-mail list, and/or downloadable to the voice mail feature of mobile devices 15 such as cell phones and/or transmittable to a hand held device such as the Palm Pilot®.

Figure 2 is a diagram of the processor elements. A profile generator 50 generates and stores a profile of the user's known interests, which includes trigger information or keywords of interest. This is accomplished for example through user input, by having the user respond to a series of queries, by creating a default profile based on user characteristics which are modified by the user, or by monitoring user activity to discover areas of interest. The rule generator 52 generates the association rules which logically combine each trigger with a variety of contexts to determine which supplementary information should be displayed to the user. The recognition engine 54 compares each trigger with the transcript text and determines whether the trigger exists as a keyword in the text. When a trigger is matched, the retrieving portion 56 retrieves the supplementary information and the formatting portion 58, formats the data for display. The context monitor 60, monitors the context to see whether it is changing due to the display of a new program segment. When a context change occurs, the context monitor 60 accesses the secondary storage 18 to retrieve a new subset of association rules.

The data updater 62 is used to update the supplementary information to incorporate new web sites, for example, or to reflect the results of searches performed by various search engines. The repetition counter 64 counts the frequency with which a particular piece of information is requested and the clickstream monitor 66 measures the frequency with which a user requests supplementary data in general. These intelligent agents

work in conjunction with the retrieval modifier 68 to modify the type of information and amount of information presented to the user.

Figures 3a and 3b are flow diagrams illustrating the method of the invention. To begin, in step S201, the input video is input to a receiver. The video is in analog or digital form. The transcript extractor, which is separate from or incorporated into the processor, extracts the transcript text in step S202 and identifies the beginning and end of each video segment. Next, in step S203, the processor retrieves the keywords from the transcript text. Extraction of keywords is well known in the art and one such method of extraction is described in U.S. Patent No.5,809,471 to Brodsky, entitled "Retrieval of additional information not found in interactive TV or telephony signal by application using dynamically extracted vocabulary." As shown in Figure 4a, these keywords 152 are extracted from the transcript text 150 and expanded 154 to achieve more meaningful and complete results, by associating them with synonymous or related keywords as shown in Figure 3a step S204. A thesaurus is used for this purpose or a database such as Wordnet®. Wordnet® is an on-line lexical reference system whose design is inspired by current psycholinguistic theories. The various parts of speech are organized into synonym sets, each representing one underlying lexical concept.

Keywords can also be expanded by identifying the theme of the transcript text. For example, the presence of the trigger “economy” in transcript text can be derived, when a number of words such as “inflation”, "Alan Greenspan", and "unemployment rate" are simultaneously present. Similarly, the presence of the trigger "President Clinton" can be derived if the keyword "President of the United States" is present in the transcript text.

Special rules apply when the supplementary data is contained in reference tools such as dictionaries and encyclopedias, as shown in Figure 4 114 132. In one mode, triggers are mapped to a variety of keywords depending on the level of understanding of the viewer. For example, if the viewer is a child or foreign-speaking viewer, the trigger "unemployment" would be mapped to the keyword phrase "without a job" but would not be mapped to the keyword "redundancy." In an alternate mode, the keywords are expanded as described above.

Parental control is implemented below the program level at the program segment or contextual level. Therefore, parents need not worry if a commercial inappropriate for children is shown during an otherwise appropriate cartoon show, for example. The child viewer is presented with a special screen only during the commercial. This special screen may take the form of a toy advertisement instead of merely a typical blocking screen.

Blocking triggers are also expanded to enhance the effectiveness of the blocking. For example, if the parent does not want the child to see video segments related to war, the trigger "war" is mapped to keywords and phrases such as "armed conflict" and "bombing." An example of trigger expansion is shown in Figure 4a 102 156.

Returning to Figure 3a, in step S205, the personal profile containing the triggers is read. The processor matches the keywords developed from the transcript text with the triggers contained in the user profile in step S206. If there is no match, the processor continues by extracting additional transcript text.

If there is a match, in step S207 of Fig. 3b, the context of the ongoing video program is identified. This is done in several ways, using either the closed caption data, EPG data, object tracking methods, or low-level feature extraction such as color, motion, texture, or shape. The context of the program segment is also extracted from the transcript text using natural language techniques. For example, Microsoft Corporation has developed software that learns by analyzing existing texts, including online dictionaries and encyclopedias, and automatically acquiring knowledge from this analysis. This knowledge is then used to help constrain the interpretation of the word "plane" in a sentence like "Flying planes can be dangerous" and to determine that the sentence pertains to aviation rather than woodworking.

Software also operates at the discourse level, using discourse analysis to identify the structure of the closed caption text and thereby its context. For example, a news program is identified because it would generally report the most important facts, "who, what, when, where, how" in its beginning. Accordingly, a program that began with the sentence "Clint Eastwood was in a gun fight, in Carmel California, at seven a.m. on Main Street, by a bystander with a home video camera" is identified as a news story. The context is also available in the EPG data from the genre and sub genre fields or a combination of fields as explained above.

Next, in step S208, the association rules are read. The association rules determine which supplementary data from a stored database should be retrieved, based upon the keyword and context. In step S209, the customized display modules are read. These modules enable the user to restrict the types of information, and therefore also the amount of information, the user wants to view. For example, the user may only wish to see the Uniform Resource Locator (URL) of a WWW page, only larger titles from the page, a page summary, or a full page. The user can choose the supplementary sources he wants to view and prioritize these sources.

In step S210, the supplementary data is retrieved from a database stored in memory. The database contains items of interest or pointers to items of interest, ancillary to the trigger. For example, the database contains any of the following: names of celebrities and public figures, geographic information such as countries, capitals, and presidents, product and brand names, assorted categories and topics.

The database is maintained and refreshed from an established set of sources. These include for example, the Bloomberg site, encyclopedias, thesauri, dictionaries, and a set of web sites or search engines. Information from the EPG and closed caption data is also incorporated into the database.

A set of refresh and cleanup rules, as shown in Figures 5 and 6 is also stored in a database or a viewer's profile, for example, and maintained for managing the size of the database or profile and its currency. For example, "stale" items such as election results and links to information about polls and the candidates would be deleted after an election takes place.

Returning to Figure 3b, in step S211, the supplementary information is formatted for display. The information is displayed in a window or superimposed unobtrusively over the main video segment. Alternatively, the information is formatted for transmittal, for example to a hand-held device such as the Palm Pilot™ distributed by Palm, Inc. or to an email account.

Figure 4 illustrates the set of association rules 100 for several triggers 102. In the table, the first column represents the triggers 102 and columns 2-4 represent the possible contexts 104, 106, 108, 110 for the example triggers shown. Beginning with the association rule 120 for the first trigger 102, "Clint Eastwood", when this trigger 102 appears in a user's profile, one of three different items of supplementary information 116, 118, 120 are retrieved for display, depending on the context in which Clint Eastwood appears in the video segment being viewed. Although only one link is shown in each box of the example table, multiple links can exist. If Clint Eastwood appears in a commercial, the system will link to the WWW page located at www.imdb.com and display the page in accordance with the customized display model. If Clint Eastwood appears on a talk show, the talk show segment where he appears will be stored for retrieval 118 and/or an alert sent to the viewer in real-time. Alternatively, an offline alert is transmitted for later viewing, notifying the viewer that the segment has been stored.

Alerts are automatically or manually retrieved. Alert transmission is also keyed to a topic such that the alert is displayed the next time a Clint Eastwood movie is

shown. If Clint Eastwood appears on a news program, the system will link to the WWW page located at www.cnn.com. Alerts have priorities enabling the user to select the circumstances when the user wants to be notified. For example, a user may only want to view alerts pertaining to severe weather warnings.

The second association rule 122 for the trigger 102 Macedonia deals with 4 different contexts. If the trigger "Macedonia" appears in an advertisement, the system links to the WWW page at www.travel.com 130. If Macedonia is the subject of a talk show, the system links to an entry for "Macedonia" in Compton's Encyclopedia 132. If Macedonia is the subject of a news show, the user is tuned to the station where the program is being aired 134. If Macedonia is the subject of a program sponsored by National Geographic magazine, the system links to www.yahoo.com/maps 136 to display a map of Macedonia.

Association rules 3-5 124 126 128 should be interpreted in the same manner as the above examples. As shown in the table, when certain triggers 102 such as "Meryl Streep" appear in transcript text, the system will only provide supplementary information for certain contexts. In the case of "Meryl Streep", supplementary information is only supplied for the Talk Show and News contexts. If desired, such a rule is broadened to apply to a list of well-known actors or all actors.

Figure 4a illustrates how both the triggers and keywords can be expanded to retrieve supplementary information. For the example transcript text 150 shown, the keyword 152 "Lyme Disease" is extracted from the transcript text 150. The keyword 152 is then expanded to map to the additional key words "tick", "tick bite", "bull's eye rash" and "deer tick." If any of these expanded keywords appear in the transcript text, supplementary information related to Lyme Disease will be retrieved.

Figure 4a also illustrates how triggers are expanded. The trigger 102 "Lyme Disease" is expanded 156 to include the related terms "tick bite", "West Nile virus, and "mosquito spraying." Accordingly, if the transcript text 150 contains any of the expanded triggers the segment is stored, for example.

Figure 5 illustrates how a learning model is implemented to continually update the customized display modules and association rules. The repetition counter 20 maintains a count of how often the user requests the same supplementary data, for example by clicking on a URL. Also, more than one piece of supplementary information may be retrieved bv the retrieving portion 56 of the processor, shown in Figure 2, for each segment and the user may select the information the user wishes to view. If a user requests a particular piece of supplementary data less than a predetermined amount of times, the stored association rules 26

are updated by the retrieval modifier 24 such that the supplementary data is eliminated from the rule or the rule is modified to include a new source. The clickstream monitor 22 monitors how frequently the user requests any supplementary data. If the user selects supplementary data less than a predetermined amount of times, the custom display module 28 for that user is modified by the retrieval modifier 24 such that less information is presented to the user.

Figure 6 illustrates how the dynamic association rules database is updated and maintained. The database contains items of interest or pointers to items of interest that can provide ancillary information, when triggered by a match between a keyword in the transcript text and a trigger in the user's profile. The database is updated over time to reflect current events and to match the evolving user profile.

The existing data sources set 36, specifies the data sources from which the association rules database 26 is constructed. The data sources set 36 which includes both external data 38 from a variety of published sources, proprietary information, and data from the Internet 14 is updated by the data updater 40 to incorporate new web sites, for example, or to reflect the results of searches performed by various search engines. A set of refresh rules 32 is maintained to keep the size of the database at a preset limit. According to a set of established priorities, information is deleted when necessary. A set of cleanup rules 34 is also maintained which specify when and how "stale" information can be deleted. Information in certain categories is date stamped, and information older than a preset number of months and/or years is deleted.

Figure 7 illustrates an embodiment in which the supplementary information 70 is displayed superimposed unobtrusively over the main video segment. The supplementary information appears at the bottom of the picture.

Figure 8 illustrates an embodiment in which a set-top box 75 comprises a receiver 2, which receives video program and transcript text. A transcript text extractor and segmenter 4 extracts the transcript text 150 from the video signal and associates it with segments of the video program such as commercials and news flashes. A processor system 6 includes processing elements well known in the art -- an input/output portion 8, a memory 10, and a processor 12. Via a communication means 17, the processor system retrieves information supplemental to the video program from a variety of sources. Three of these sources, the Internet 14, proprietary (non-public) databases 13, and mobile devices 15 such as PDAs are shown in the figure as examples. The communication means 17 can connect to other devices not specifically shown, via wireless means, cable modem, a digital subscriber line, or a network, for example. The secondary storage 18 is used to store the supplementary

information as well as the rules for retrieving the information. The set-top box can be interfaced to a display such as a PC display or a television.

Figure 9 illustrates another embodiment in which a television 80 comprises a receiver 2, a transcript text extractor and segmenter 4, a processor system 6, secondary storage 18, a communication means 17, and a display 16. The processor system 6 includes processing elements well known in the art - - an input/output portion 8, a memory 10, and a processor 12. The television 80 interfaces to sources of supplementary information via the communication means 17 which interfaces to the Internet 14, proprietary sources 13 and mobile devices 15, for example.

The present invention has been described with respect to particular illustrative embodiments. It is to be understood that the invention is not limited to the above-described embodiments and modifications thereto, and that various changes and modifications may be made by those of ordinary skill in the art without departing from the spirit and scope of the appended claims.

CLAIMS:

1. An association method for retrieving information supplemental to a video program comprising the steps of:

receiving the video program (2);

identifying in the video program at least one segment (4);

receiving classification data for said at least one segment (4,2);

receiving transcript text for the video program (4);

identifying a user profile for a video program viewer (50);

identifying a set of rules (52) incorporating the classification data, for associating the supplementary information with the video program, when the transcript text and the user profile satisfy a set of conditions; and

automatically retrieving the supplementary information based upon the set of rules for display on a display (56).

2. The method according to Claim 1, wherein the set of rules (100) includes information from the user profile (102).

3. The method according to Claim 2, wherein the user profile contains at least one trigger (102) which identifies a topic of interest to the video program viewer.

4. A method according to Claim 3, wherein the set of conditions specifies that a recognition engine (54) retrieve the supplementary information only when a keyword in the transcript text matches (S206) the at least one trigger (102) in the user profile.

5. The method according to Claim 1, wherein the transcript text is comprised of closed caption text, video text, program transcripts or electronic program guide information.

6. The method according to Claim 1, wherein the transcript text (150) is generated by speech recognition software.

7. The method according to Claim 1, further including the step of receiving at least a portion of the set of rules (100) from a mobile device (15) or a third-party source (13).

8. The method according to Claim 1, wherein at least part of the supplementary information and pointers to the supplementary information are stored in a database (26) or transmitted to a personal digital assistant (15) or to an electronic mail address (14).

9. The method according to Claim 1 wherein the retrieval of the supplementary information (116,118,120) is in real-time.

10. The method according to Claim 1, wherein the supplementary information (116,118,120) is formatted for display in a window (70) or for superimposition over the video program on a display (16).

11. The method according to Claim 1, wherein the supplementary information is text information (114) or a page from the World Wide Web (116).

12. The method according to Claim 5, further including the step of automatically selecting the set of rules (100) for each video program segment from the electronic program guide information (150).

13. The method according to Claim 3, further including the step of automatically selecting the set of rules (100) by applying natural language processing to the transcript text (150) for each video program segment to identify whether a keyword (S203) in the transcript text (4) matches a trigger (102) in the user profile.

14. The method according to Claim 3, further including the step of identifying at least one keyword (S203, 152) in the transcript text (150), expanding the at least one keyword (S204, 152) to include related terms (154), and retrieving the supplementary information (S210) when the keyword or related terms matches (S206) the at least one trigger (102) in the user profile.

15. The method according to Claim 3, further including the step of automatically generating the set of rules (52) by applying discourse analysis to the transcript text (150) for

each video program segment to identify whether a keyword (152) in the transcript text (150) matches a trigger (S206,102) in the user profile.

16. The method according to Claim 3, further including the step of expanding at least one trigger (154) in the user profile to include related terms, identifying at least one keyword in the transcript text, and retrieving the supplementary information when the trigger or related terms matches the at least one keyword in the transcript text.

17. The method according to Claim 8, further including the step of deleting (40) supplementary information (26) or pointers to supplementary information added to the database before a certain date or related to events that have terminated.

18. The method according to Claim 11, wherein only the Uniform Resource Locator (URL) (28,70) of the page or wherein a portion of the page (28) which is less than the entire page or wherein a summary of the page (28) is displayed.

19. The method according to Claim 1, further including the step of monitoring (22) the amount of supplementary information viewed by the video program viewer, and the frequency (20) with which the video program viewer views the supplementary information, and varying (24) the amount of supplementary information formatted for display correspondingly, according to a predetermined formula.

20. The method according to Claim 1, wherein the supplemental information is included in an electronic mail message (15) or is downloaded (17) to a personal information manager (15).

21. An apparatus for retrieving information supplementary to a video program, the apparatus comprising:

a receiver (2) which receives the video program, classification data for the video program, and transcript text for the video program;

a transcript extractor (4) which identifies at least one segment within the video program and associates transcript text with said one segment;

a context monitor (60,S207), which monitors the classification data (104,106,108,110) for each segment thereby identifying a context for each segment;

a profile generator (50), which establishes a user profile for a video program viewer;

a rule generator (52), incorporating the classification data (102,104,106,108,110), which establishes a set of rules (100) for associating supplementary information (116,118,120) with the video program, when the transcript text (150) and the user profile (102) satisfy a set of conditions;

a retrieving portion (56), which retrieves the supplementary information (116,118,120), based upon the set of rules (100);

a formatting portion (58) which formats (S211) the retrieved supplementary information for display along with the video program.

22. An apparatus according to Claim 21 wherein the retrieving portion retrieves (S210) the supplementary information (116,118,120) when a trigger (102) within the user profile matches (S206) a keyword (152) within the transcript text.

23. An apparatus according to Claim 22, wherein at least one trigger (102) in the user profile is expanded (156) to include related terms and the trigger and the related terms are compared (S206) with the keyword (152).

24. An apparatus according to Claim 22, wherein at least one keyword (152) within the transcript text (150) is expanded (154, S204) to include related terms and the trigger (102) is compared with the keyword (154) and the related terms.

25. An apparatus according to Claim 21, wherein the retrieving (S207, 104,106,108,110) portion (56) retrieves information for the segment based upon the context of the segment.

26. Computer-executable process steps to retrieve information supplemental to a video program, the computer-executable process steps being stored on a computer-readable medium (18) and comprising:

a receiving step (S201) to receive the video program, classification data describing the video program, and transcript text for the video program;

a context identifying step (S207) to identify at least one segment in the video program and the context of the segment based upon the classification data;

a keyword identification step (S203) to identify keywords in the transcript text for the at least one segment in the video program;

a keyword expanding step (S204) to expand the keywords to include related terms;

a personal profile retrieving step (S205) to retrieve a user profile for a viewer viewing the video program;

a keyword matching step (S206) to match the keywords and the related terms with the at least one trigger in the user profile;

an association rules retrieving step (S208) to retrieve a set of rules specifying which information supplemental to the video program will be retrieved, depending upon the identified context;

a retrieving step (S210) to retrieve the supplementary information based upon the set of rules when the keyword matching step is successful; and

a formatting step (S211) to format the retrieved supplementary information for display;

27. A signal (160), embodied in a carrier wave, representing a video program (162) and information supplemental thereto (116,118,120), comprising video program classification data (104,106,108,110); transcript text (150); a user profile (102); and rules (100) incorporating the video program classification data, for associating the supplementary information with the video program when the transcript text and the user profile satisfy a set of conditions (S206).

28. An apparatus for retrieving and displaying information supplemental to a video program comprising:

means (2) for receiving the video program (162);

means for identifying in the video program at least one segment (4);

means for receiving program classification data describing the at least one segment (4,2);

means for receiving transcript text (150) for the video program and associating the transcript text with the at least one segment (4);

means for retrieving a user profile for a video program viewer (50);

means for identifying (52) a set of rules (100), incorporating the classification data (104,106,108,110), for associating the supplementary information (116,118,120) with

the video program, when the transcript text and the user profile (102) satisfy a set of conditions (S206);

means for retrieving the supplementary information based upon the set of rules (56,S210); and

means for formatting (58) the supplementary information for display along with the video program.

29. A set-top box (75) for a video program viewer, comprising:

receiving means (2) which receives a video program (102), classification data for the video program (104,106,108,110), and transcript text (150) for the video program;

transcript text extraction and segmenting means (4) which identifies at least one segment in the video program and associates transcript text with the at least one segment;

communication means (17) which connects to at least one information source (14,13,15) and receives information supplemental to the video program (116,118,120);

processor means (6) which

a) retrieves a user profile (50) for the video program viewer which contains at least one trigger (102) reflecting an interest of the video program viewer,

b) associates the classification data with the at least one segment (60, S207),

c) identifiifes a set of rules (52) incorporating the classification data, for associating the supplemental information with the segment,

d) searches the transcript text for a trigger contained in the user profile (54),

e) retrieves the supplemental information (56), using the communication means (17) and based upon the set of rules (100), when the trigger (102) is contained within the transcript text (150), and

f) formats (58) the retrieved supplemental information for display; and

storage means (18) which stores the transcript text, the user profile, the set of rules, and the supplemental information.

30. The set-top box (75) according to Claim 29, wherein the receiving means receives a digital video program.

31. The set-top box according to Claim 29 (75), wherein the processor (12) decodes and formats the digital video program for display on an analog display.

32. The set-top box (75) according to Claim 29, wherein the video program viewer selects a destination (15) where the supplementary information will be transmitted via the communication means (17).

33. The set-top box (75) according to Claim 29, wherein more than one type of supplementary information (116,118,120) is retrieved by the processor (12) for each segment, the retrieved supplementary information is automatically placed in an order of priority according to the user profile (S209), and the supplementary information with highest priority is formatted for display (S211) by default.

34. The set-top box (75) according to Claim 29, wherein more than one type of supplementary information (116,118,120) is retrieved by the processor (12) for each segment, and the video program viewer selects the retrieved supplementary information the video program viewer wishes to view.

35. A television set (80) comprising:

receiving means (2) which receives a video program (162), classification data for the video program (104,106,108,110), and transcript text (150) for the video program;

transcript text extraction and segmenting means (4) which identifies at least one segment in the video program and associates transcript text with the at least one segment;

communication means (17) which connects to at least one information source and receives information supplemental to the video program;

processor means (12) which

a) retrieves a user profile (50) for a video program viewer which contains at least one trigger reflecting an interest of the video program viewer,

b) associates the classification data with the at least one segment (4,2),

c) identifies a set (52) of rules (100), incorporating the classification data, for associating the supplemental information with the segment,

d) searches the transcript text (54) for a trigger (102) contained in the user profile,

e) retrieves the supplemental information (116,118,120), using the communication means (17), and based upon the set of rules (100), when the trigger (102) is contained within the transcript text, and

f) formats (58) the retrieved supplemental information for display;

storage means (18) which stores the transcript text, the user profile, the set of rules, and the supplemental information; and

display means which displays the video program and the retrieved and formatted supplemental information.

36. Computer-executable process steps to retrieve information supplemental to a video program, the computer-executable process steps being stored on a computer-readable medium (18) and comprising:

a receiving step (S201) for receiving the video program, classification data describing the video program and transcript data for the video program;

a segmenting step (S202) for identifying at least one segment in the video program and classification data for the segment;

a first identifying step (S205) for identifying a user profile for a video program viewer;

a second identifying step (S208) for identifying a set of rules incorporating the classification data, for associating the supplementary information with the video program, when the transcript text and the user profile satisfy a set of conditions; and

a retrieving step (S210) for automatically retrieving the supplementary information based upon the set of rules.

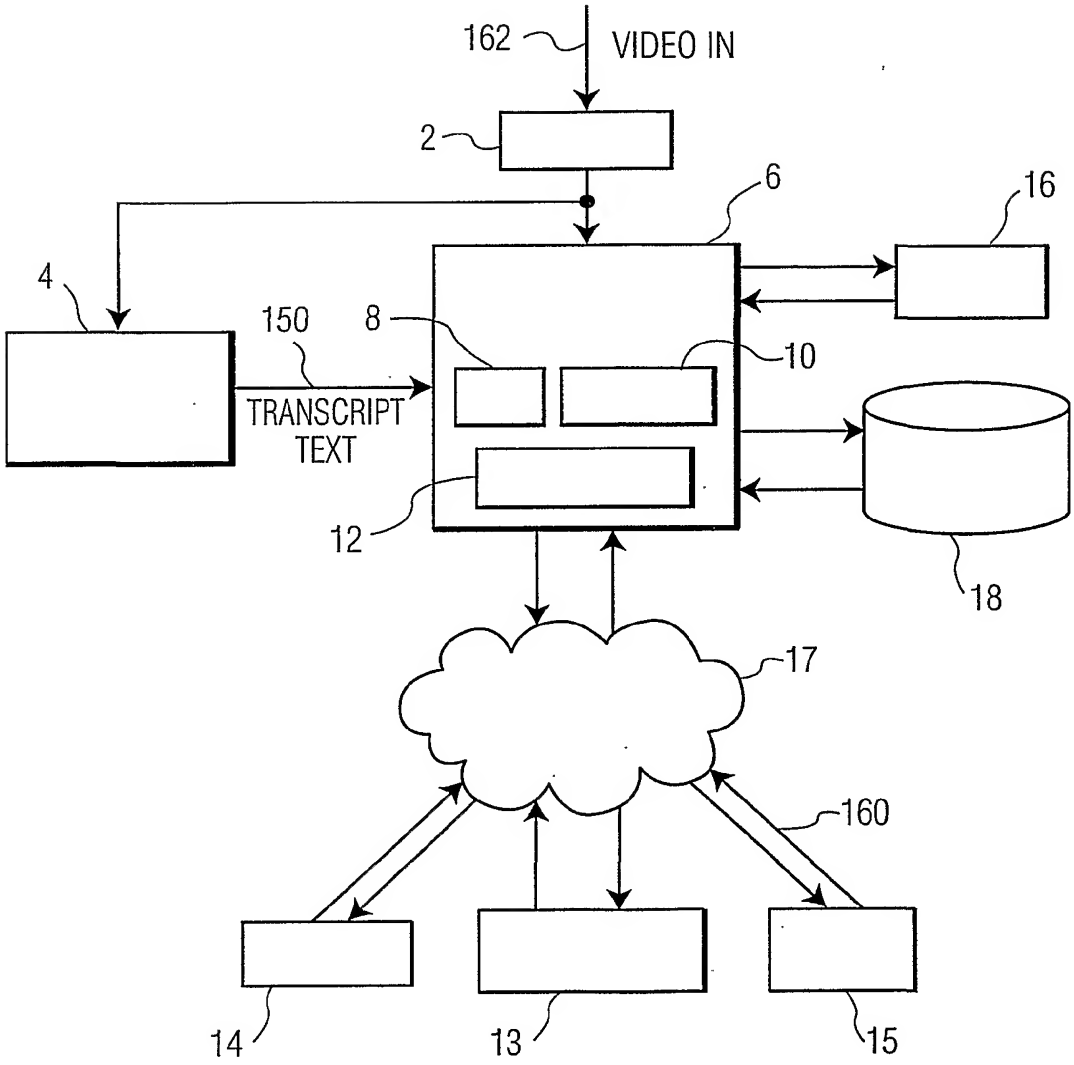
162
VIDEO IN
2
6
16
4
150
8
10
TRANSCRIPT
TEXT
12
18
17
160
14
13
15

FIG. 1

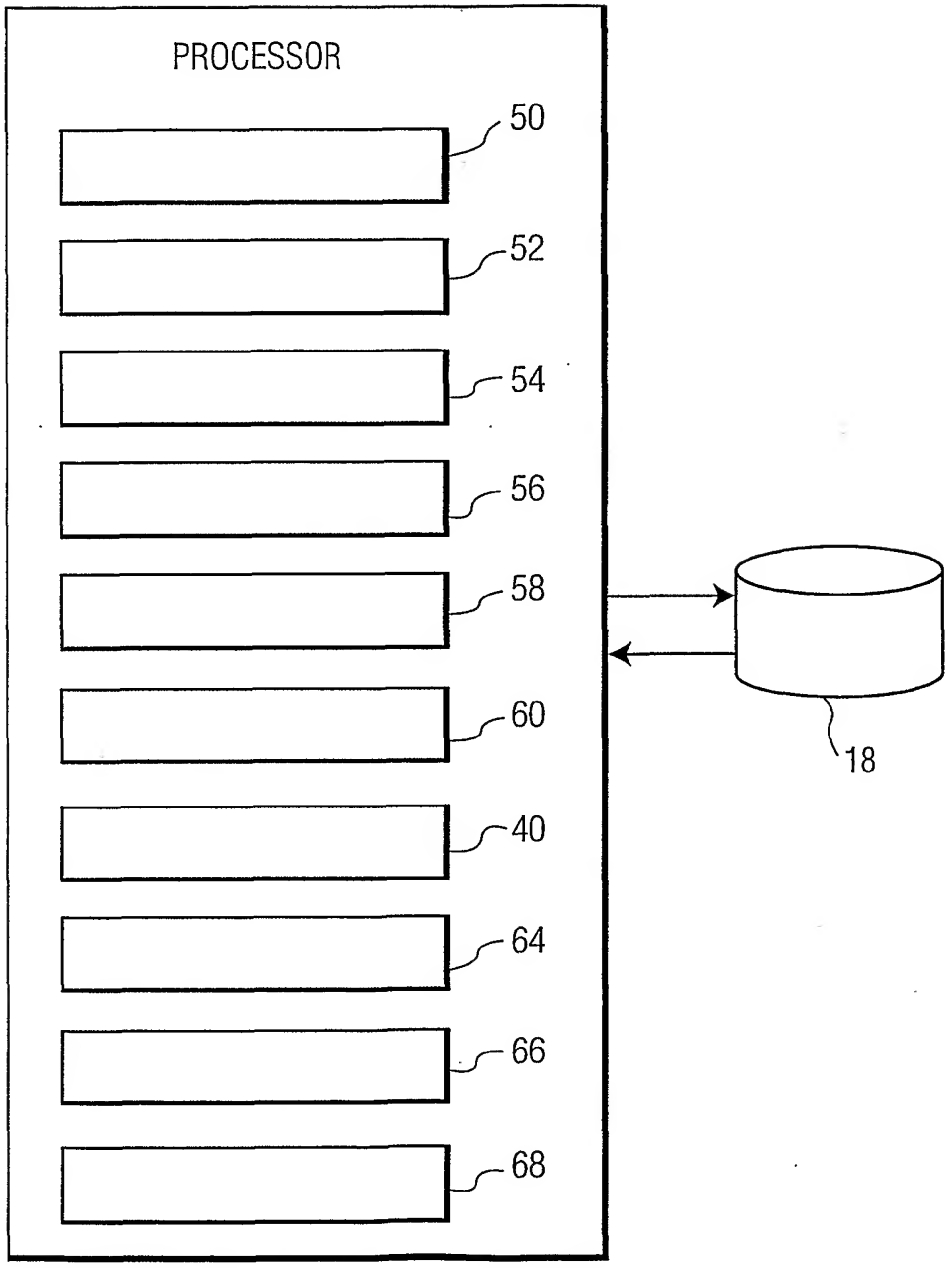

FIG. 2

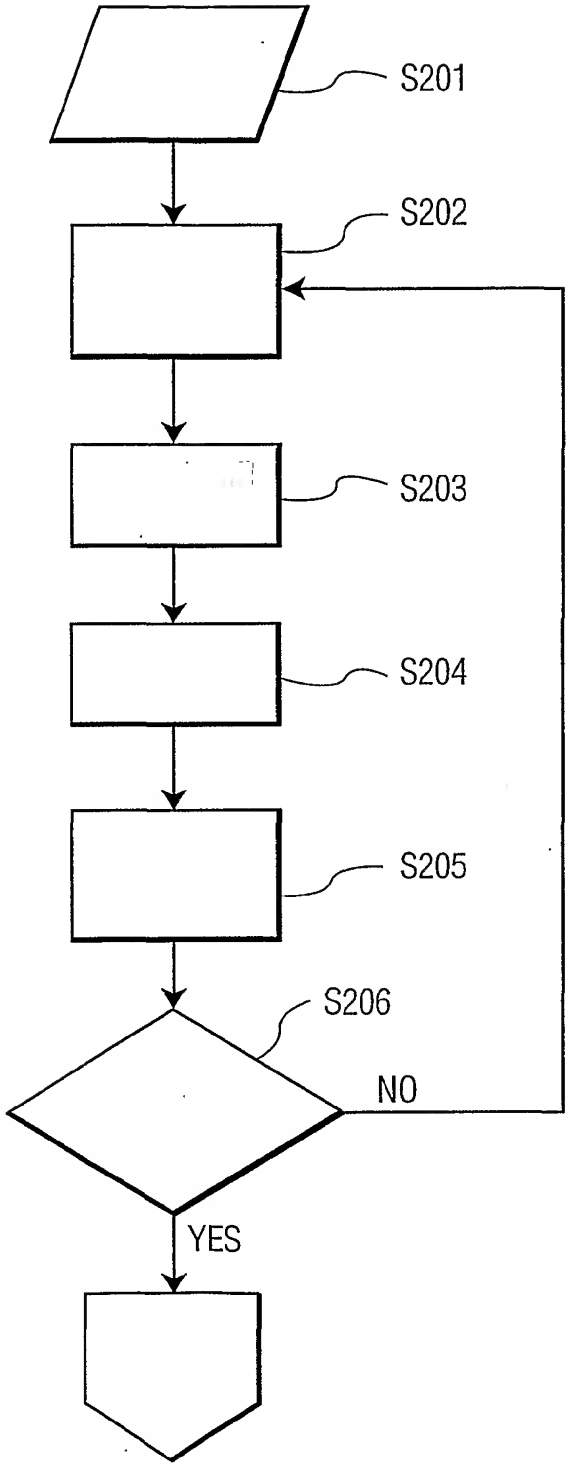

FIG. 3a

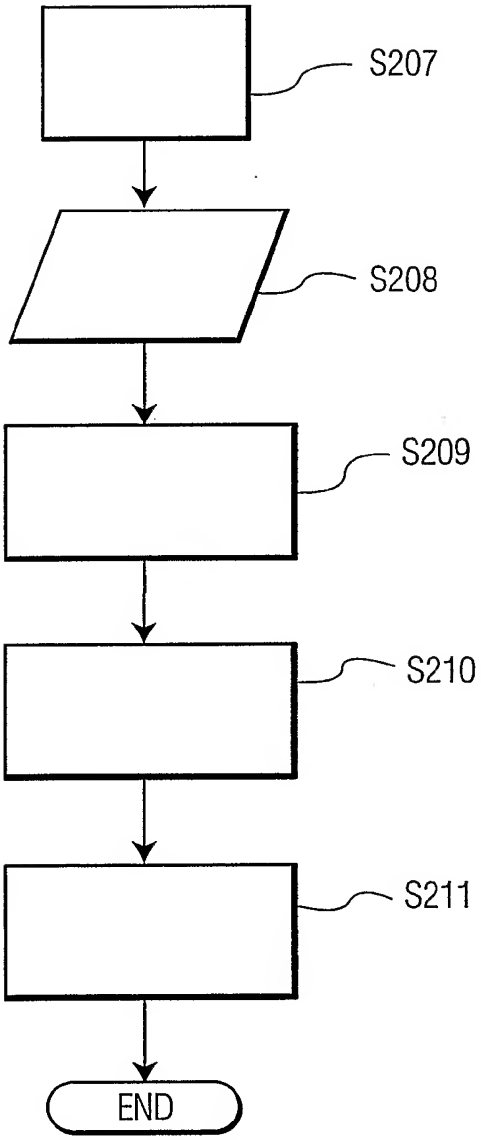

FIG. 3b

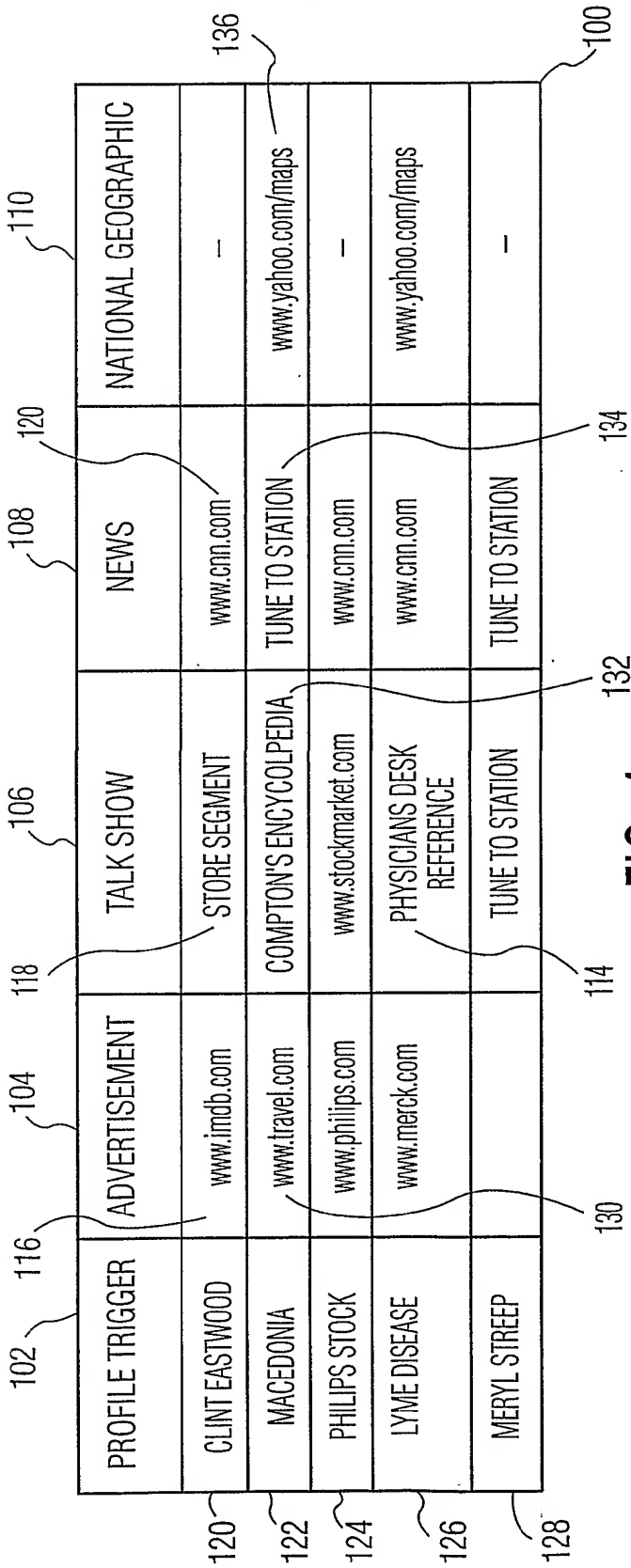
PROFILE TRIGGER
ADVERTISEMENT
TALK SHOW
NEWS
NATIONAL GEOGRAPHIC
CLINT EASTWOOD
www.imdb.com
STORE SEGMENT
www.cnn.com
–
MACEDONIA
www.travel.com
COMPTON'S ENCYCOLPEDIA
TUNE TO STATION
www.yahoo.com/maps
PHILIPS STOCK
www.philips.com
www.stockmarket.com
www.cnn.com
–
LYME DISEASE
www.merck.com
PHYSICIAN'S DESK REFERENCE
www.cnn.com
www.yahoo.com/maps
MERYL STREEP
TUNE TO STATION
TUNE TO STATION
–
102
116
104
118
106
108
120
110
136
100
120
122
124
126
128
130
114
132
134

FIG. 4

TRANSCRIPT TEXT 150

COMING UP NEXT IS A DISCUSSION OF LYME DISEASE WITH DR. JOHN JONES

KEYWORD 152

LYME DISEASE

KEYWORD EXPANSION 154

LYME DISEASE
TICK
TICK BITE
BULL'S EYE RASH
DEER TICK

TRIGGER 102

LYME DISEASE

TRIGGER EXPANSION 156

LYME DISEASE
TICK BITE
WEST NILE VIRUS
MOSQUITO SPRAYING

FIG. 4a

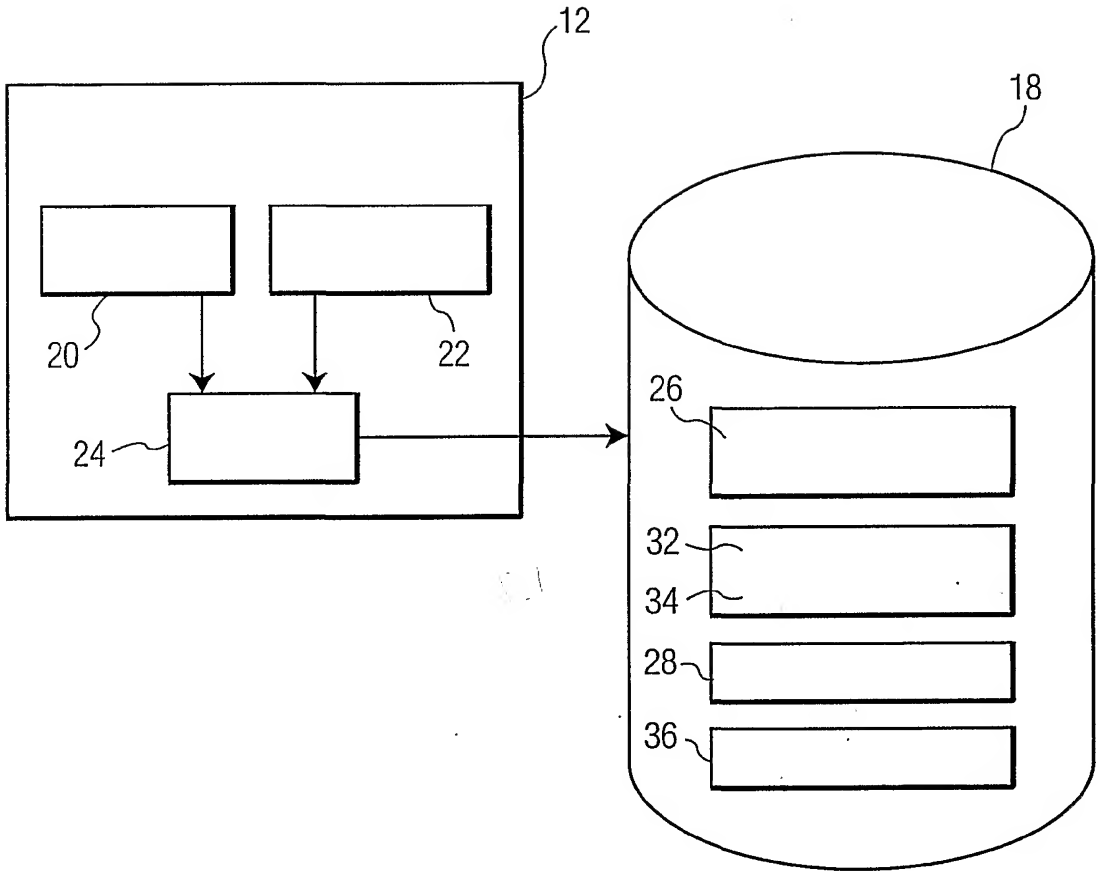

FIG. 5

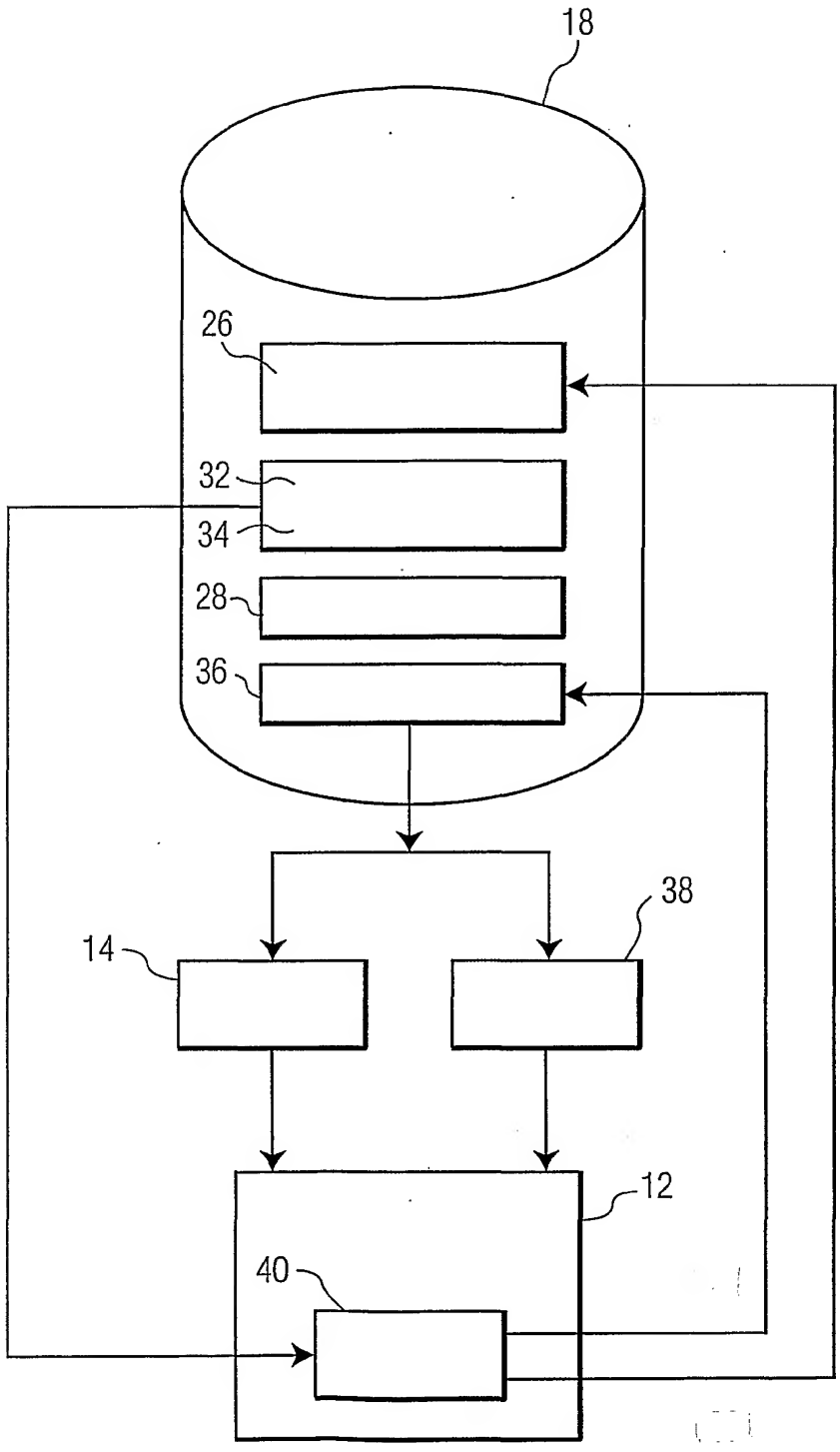

FIG. 6

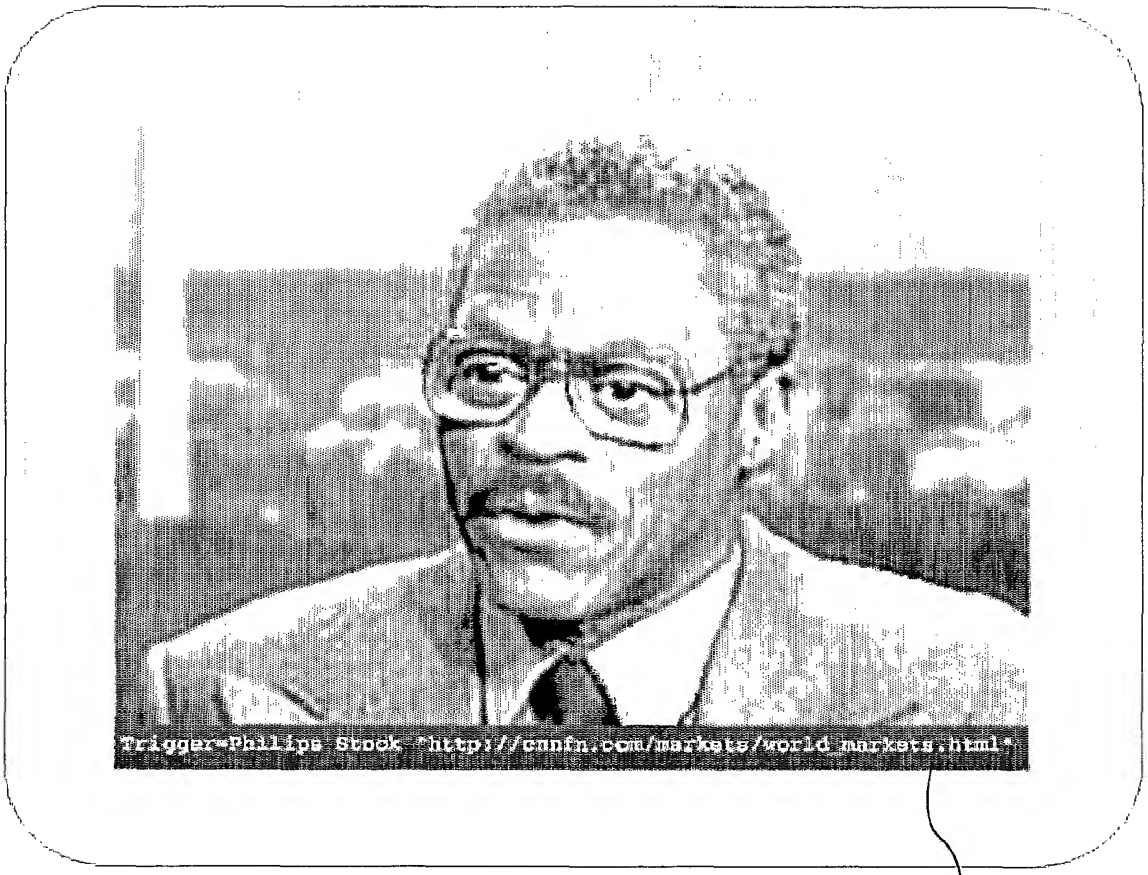
Trigger="Philips Stock" "http://cnnfn.com/markets/world_markets.html"
70

FIG. 7

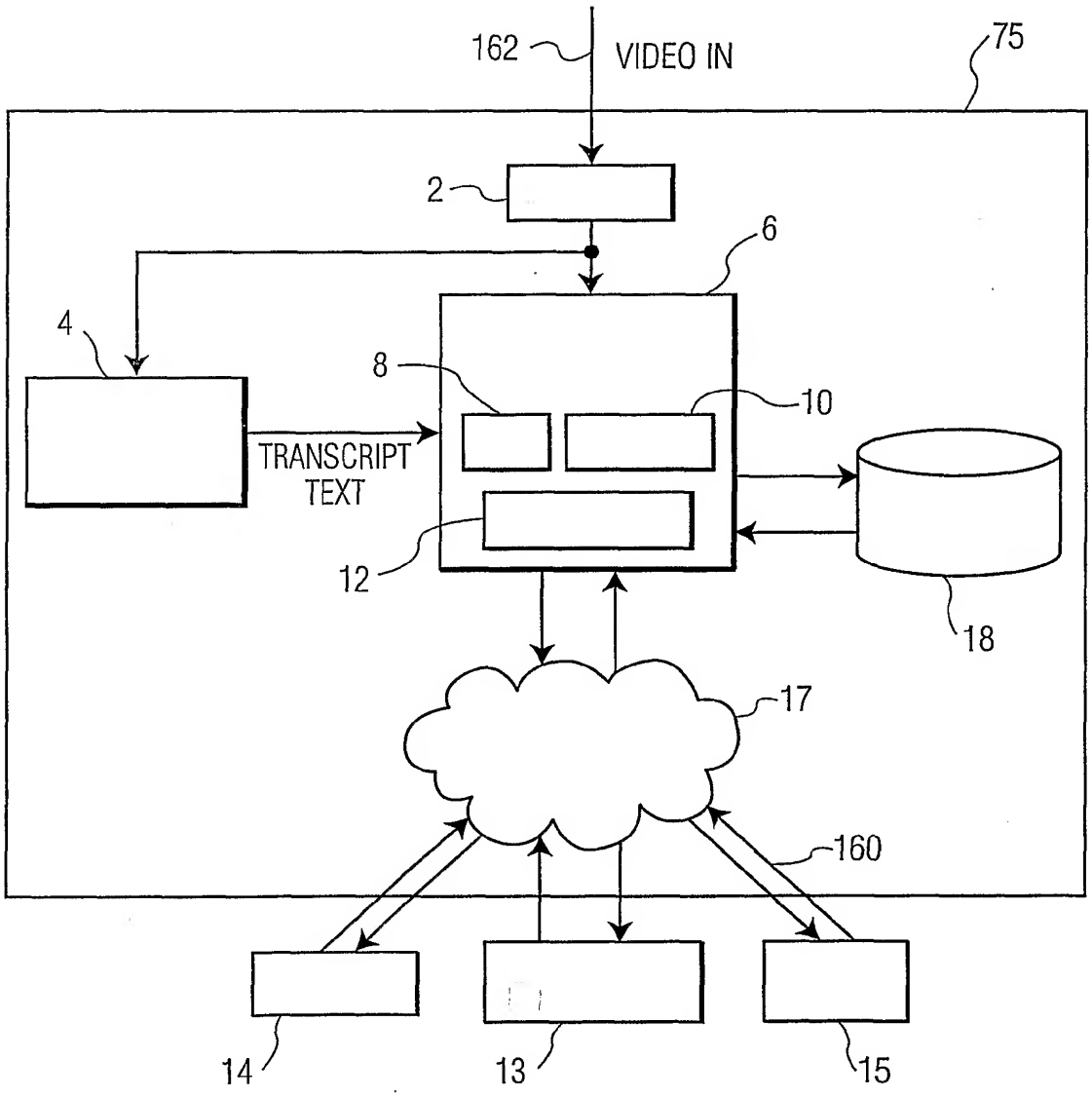

FIG. 8

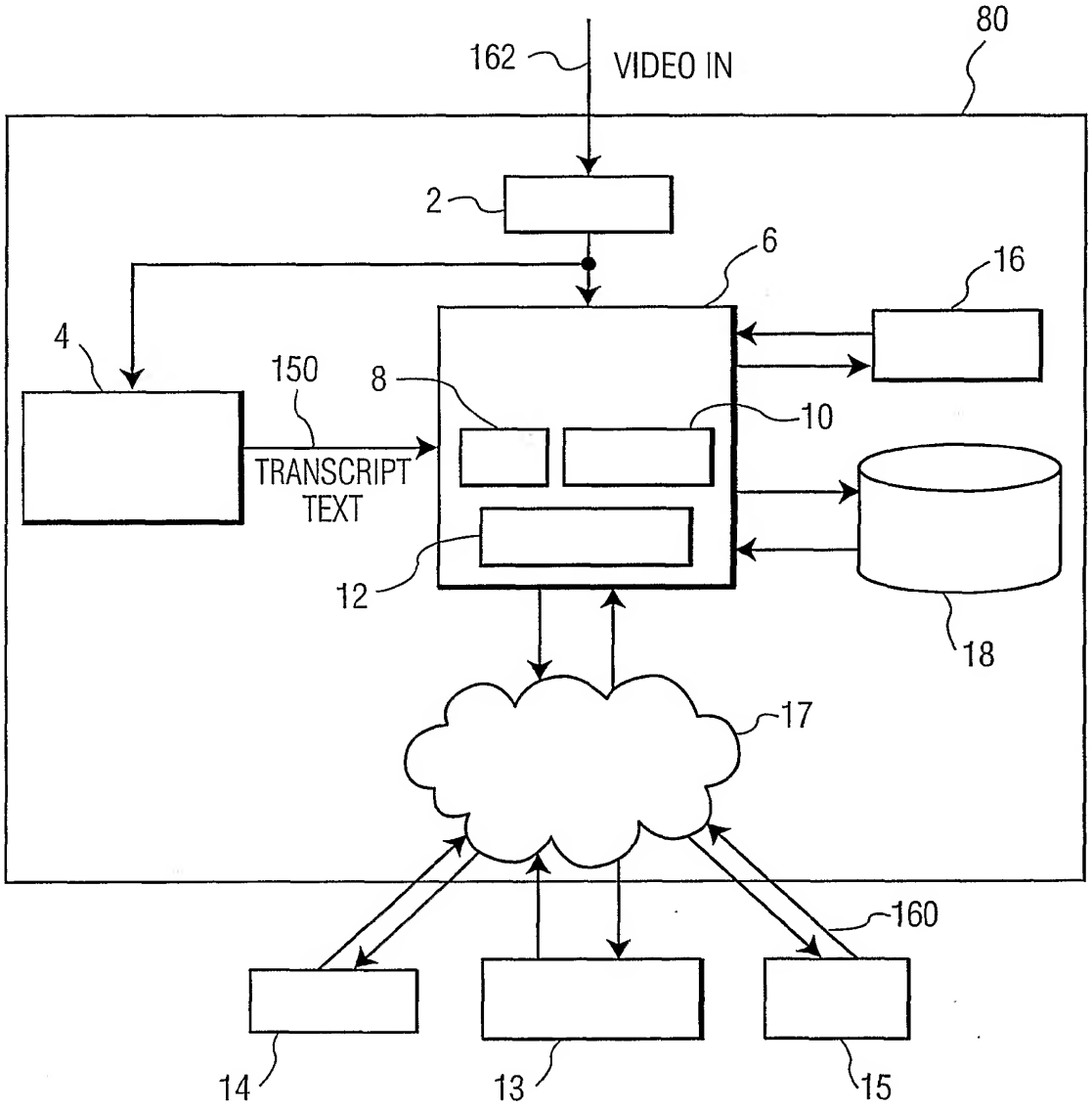
80
162
VIDEO IN
2
6
16
4
150
8
10
TRANSCRIPT
TEXT
12
18
17
160
14
13
15

FIG. 9